

# SYSTEM AND METHOD FOR STREAMING MEDIA SERVER SINGLE FRAME FAILOVER

## FIELD OF THE INVENTION

**[0001]** This invention relates to the field of fault-tolerant systems and, in particular, to the field of fault-tolerant digital media systems.

## BACKGROUND OF THE INVENTION

**[0002]** As the number of subscribers for video-on-demand (VOD) services continues to grow, consistent and dependable delivery of such services becomes critical. One common mechanism for ensuring that data transmissions are not interrupted by server failure is known in the art as "failover." Failover refers generally to the technique of automatically switching to a backup server when a primary server fails. Failover is an important fault-tolerance feature of systems that must be constantly available such as email systems and database servers.

**[0003]** Systems with failover capability typically group servers into failover pairs. Each failover pair includes a primary server that is active and a secondary server that is brought online only when a failover occurs. Other designations used interchangeably to refer to the primary and secondary servers are master and slave, active and standby, and primary and backup.

**[0004]** Although vendors often assert that their failover systems provide "instant" or "immediate" failover, these systems typically initiate a failover seconds or even minutes after a failure occurs. For many types of applications this delay is acceptable. For example, a delay of several seconds will typically go unnoticed by users accessing a database or email server due to the non-time-sensitive nature of the data being accessed. But such delays cannot be tolerated by providers of premium digital-media services.

**[0005]** For example, suppose a customer accesses a VOD service provided by a local cable television provider. The cable television provider will typically employ a plurality of digital media servers to provide VOD services to its customers. If any VOD server currently delivering content experiences a catastrophic failure and goes offline, the movie is disrupted and the customer's movie viewing experience is unsatisfactory.

[0006] Moreover, due to the enormous storage requirements and time-sensitive nature of delivering digital media, the process of failover for a digital media server is far more complex to implement than that of a database or Web server. The growth of VOD, subscription VOD (SVOD), and everything-on-demand (EOD) services, combined with the growing subscriber density served by each digital media server, requires stricter fault tolerance levels for asset availability.

## SUMMARY OF THE INVENTION

[0007] A system and method for providing failover capability is disclosed. In a preferred embodiment, a plurality of digital media servers are divided into failover pairs. In each pair, one server is designated as the primary server, and one server is designated as the secondary server. The secondary server preferably maintains up-to-date asset and other information that mirrors the primary server.

[0008] The operational status of the primary server is preferably verified on a continuing basis using one or more techniques which may include local monitoring by the primary server of critical processes and remote monitoring of a periodic "heartbeat" generated by the server. The heartbeat frequency is preferably greater than the session's interval rate. If the primary server is operating correctly, it's output is streamed to the user and the second server's output is discarded. If, however, the primary server is not operating correctly, a failover is triggered and the second server takes over delivery of the data to the user.

[0009] Because the secondary server mirrors the primary server's operational state and processes asset requests in parallel with the primary server, and because the primary server is continuously monitored for failures that may affect its ability to deliver requested data to the client, the present system and method can transfer all functionality (including asset access and serving functions) from a primary server to a secondary server in less than one video frame. Consequently, the present system and method eliminates virtually all disruption of service to a VOD customer that might otherwise be experienced due to server failure.

[0010] In one aspect, the present invention is directed to a method for data delivery comprising a first server computer connected to a first network, a second server computer connected to the first network, said first and second servers being interconnected via a second

network, the method comprising:

- synchronizing parameters of the first and second server computers;
- receiving an asset request from a user via the first network;
- processing the asset request by the first and second server computers;
- determining the operational status of the first server computer, wherein if a failure is not detected, transmitting the asset by the first server via the first network,
- if a failure is detected, transmitting the asset by the second server via the first network.

[0011] In another aspect of the present invention, the method further comprises the steps of detecting a failure and transmitting the asset by the second server computer via the first network being performed within one interval.

[0012] In another aspect of the present invention, the method further comprises the interval being one video frame in duration.

[0013] In another aspect of the present invention, the method further comprises the second server computer initiating data synchronization.

[0014] In another aspect of the present invention, the method further comprises the first server computer initiating data synchronization.

[0015] In another aspect of the present invention, the method further comprises a synchronization component initiating data synchronization.

[0016] In another aspect of the present invention, the method further comprises wherein the step of detecting a failure comprises monitoring a plurality of signals transmitted by the first server computer during one interval.

[0017] In another aspect of the present invention, the method further comprises the plurality of signals being transmitted at a frequency greater than 1 divided by the interval.

[0018] In another aspect of the present invention, the method further comprises the interval being one video frame in duration.

[0019] In another aspect of the present invention, the method further comprises a failure being determined to have occurred when a predefined number of signals are not received.

[0020] In another aspect of the present invention, the method further comprises the step of detecting a failure being performed by the second server computer.

[0021] In another aspect of the present invention, the method further comprises the step of detecting a failure being performed by a component monitor.

[0022] In another aspect of the present invention, the method further comprises the step of detecting a failure being performed by the first server computer.

[0023] In another aspect of the present invention, the method further comprises the step of detecting a failure being performed by a kernel running on the first server computer.

[0024] In another aspect of the present invention, the method further comprises one or more applications critical to the operation of the first server computer registering with the kernel.

[0025] In another aspect of the present invention, the method further comprises a failure being determined to have occurred if the kernel recognizes one or more critical application failures.

[0026] In another aspect of the present invention, the method further comprises defining one or more failover states for a server computer.

[0027] In another aspect of the present invention, the method further comprises the failover state comprising a Primary state.

[0028] In another aspect of the present invention, the method further comprises the failover state comprising a Primary\_offline state.

[0029] In another aspect of the present invention, the method further comprises the failover state comprising a Primary\_no\_secondary state.

[0030] In another aspect of the present invention, the method further comprises the failover state comprising a Failed state.

[0031] In another aspect of the present invention, the method further comprises the failover state comprising a Secondary state.

[0032] In another aspect of the present invention, the method further comprises the failover state comprising a Secondary\_offline state.

[0033] In another aspect of the present invention, the method further comprises the failover state comprising a Secondary\_synchronizing state.

[0034] In another aspect of the present invention, the method further comprises the failover state comprising a Secondary\_synchronized state.

[0035] In another aspect of the present invention, the method further comprises the failover state comprising a Secondary\_no\_primary state.

[0036] In another aspect, the present invention is directed to a method for data delivery comprising a first server operating on a first computer, a second server operating on the first computer, said first and second servers connected to a first network, the method comprising:

- storing identical data on the first and second servers;
- receiving an asset request from a user via the first network;
- processing the asset request by the first and second server;
- determining the operational status of the first server, wherein
  - if a failure is not detected, transmitting the asset by the first server via the first network,
  - if a failure is detected, transmitting the asset by the second server via the first network.

#### BRIEF DESCRIPTION OF THE DRAWINGS

[0037] Figs. 1A-B are block diagrams illustrating a preferred embodiment of a digital media server system with failover capability;

[0038] Fig. 2 is a flow chart illustrating system operation of the preferred embodiment of Figs. 1A-B;

[0039] Figs. 3A-E are composite block/flow diagrams emphasizing the timing of data delivery in a failover pair and the single-frame failover capability of the present system and method;

[0040] Fig. 4 is a block diagram illustrating a preferred network configuration of public and private interfaces for supporting a failover pair; and

[0041] Fig. 5 is a diagram illustrating a plurality of failover states and their relationship in a preferred embodiment of the present system and method.

#### DETAILED DESCRIPTION OF THE DRAWINGS

[0042] Figs. 1A-B illustrate a preferred embodiment of a digital media server system with failover capability. As shown in Fig. 1A, the system preferably comprises a failover pair 100 that includes a first server 102 and a second server 104. At any point in time, one of the servers is designated primary server 102 and the other is designated secondary server 104. As described below, both the primary and secondary servers process requests but only the primary server's output is delivered to the client unless a failover occurs.

[0043] In the example of Fig. 1A, server 102 is initially designated the primary server and server 104 is initially designated the secondary server. As described below, commands are preferably defined for use by a system administrator to set parameters in each system server that specify whether the server is a primary or secondary server, and to identify the second server in its failover pair. Alternatively, the system may be programmed to establish its own failover pairs.

[0044] In a preferred embodiment, both servers in the failover pair are capable of serving the same number of sessions, have access to the same content, and are adapted to provide the same service level. The servers' device configurations, however, need not be identical. In addition, during operation, operational parameters of both servers in the failover pair are synchronized (e.g., number of active sessions, the status of all active sessions, data port numbers, packet numbers, and packet send times). In the following description, the term "failover-sensitive parameters" is used to refer to configuration and other parameters of a server that must be synchronized with its failover partner to allow a secondary server to immediately take over streaming for a primary server if the primary server fails. Preferred embodiments of these parameters are described below. The task of synchronizing failover-

sensitive parameters may be allocated to a synchronizer component 108, or, if desired, integrated into component monitor 106 (described below) or servers 102,104.

**[0045]** In a preferred embodiment, failover pair 100 is provided with a component monitor 106 adapted to verify the operational status of components in servers 102,104. For example, component monitor 106 may be adapted to pull or push data from or to a server for purposes of evaluating the server's health. When component monitor 106 detects a failure of primary server 102, it preferably triggers a failover which transfers responsibility for delivering content to the secondary server, as described below.

**[0046]** In a preferred embodiment, a failover is triggered when primary server 102 cannot deliver a requested resource for any reason. For example, a failover may be triggered when the primary server is inoperative due to a power failure, hardware or software failure, or networking failure. When a failover is triggered, the secondary server begins delivering data and is preferably re-designated as the primary server at least until the original primary server is returned to operation or the malfunctioning primary server is replaced by a new server.

**[0047]** It should be noted that although the present description speaks primarily in terms of a single primary server and a single secondary server, the rectangle that represents a single primary server 102 in Fig. 1A may, in some embodiments, represent a first server cluster, and the rectangle that represents a single secondary server 104 in Fig. 1A may, in some embodiments, represent a second server cluster. It should also be noted that both servers in a failover pair may reside on a single physical machine, with the primary and secondary servers residing on two independent data paths.

**[0048]** As known in the art, the time required to deliver one discreet portion of a transmitted data type is referred to as an "interval." In the case of video-based media in the United States, Canada, and Japan, the National Television Standards Committee (NTSC) standard video format is approximately 30 frames per second. The transmission time for one video frame, i.e., one interval, is thus approximately 33 milliseconds. By contrast, in the case of video-based media in Europe, the Phase Alternate Line (PAL) or Sequential Couleurs a Memoire (SECAM) standard video format is approximately 25 frames per second. Thus, the transmission time for one video frame, i.e., one interval, is approximately 40 milliseconds. Although the following description refers primarily to the NTSC format, it should be

recognized that the present system and method may be applied to other formats, such as PAL, SECAM, and others, such as formats employing 24 frames per second for film-based media.

[0049] A failover in the present system may preferably be triggered in a number of ways. In one preferred embodiment, a failover is triggered if monitoring component 106 detects a failure in primary server 102. For example, component monitor 106 may be adapted to detect loss of an interface link (e.g., if a network connection is severed). When component monitor 106 detects a malfunction in primary server 102, it notifies secondary server 104 that it should begin delivering the requested data. This notification may preferably be made by component monitor 106 itself, or via a failover switch (not shown) incorporated into the component monitor. Component monitor 106 may also be adapted to monitor the status of synchronizer 108, and network connections 110, 112, 122, and 132.

[0050] Alternatively or in addition, a failover may be triggered by primary server 102 detecting a malfunction in its own operation and transmitting a failure message to secondary server 104 or component monitor 106. For example, primary server 102 may be adapted to recognize that a hardware error will prevent access to its network interface for transmitting data, and to send a message indicating this failure to either secondary server 104 or component monitor 106.

[0051] Alternatively or in addition, the primary server may be adapted to transmit a signal (referred to herein as a "heartbeat") at a predefined frequency indicating that the primary server is operating properly. A system administrator preferably defines the heartbeat frequency in milliseconds. If, for example, the defined heartbeat frequency is 5 milliseconds, a properly operating primary server will transmit approximately 6.5 heartbeats per NTSC frame. Heartbeats may be monitored by secondary server 104, component monitor 106, or other suitable monitoring components.

[0052] In a preferred embodiment, the system administrator defines the number of heartbeats that may be missed, i.e., not detected by the secondary server or component monitor, before a failure is determined to have occurred.

[0053] Fig. 1B is a block diagram illustrating the system of Fig. 1A after a failover is triggered. As shown in Fig. 1B, when a failover is triggered, secondary server 104 takes over for primary server 102 and delivers requested content to the client via a network connection 114.

**[0054]** Operation of failover pair 100 will now be further described in connection with Fig. 2. For purposes of illustration, an interval in the following description is assumed to be equal to one NTSC video frame. As shown in Fig. 2, an incoming request 200 is simultaneously routed to both the primary server and secondary server. In a preferred embodiment, network spoofing techniques are employed to make primary server 102 and secondary server 104 appear as a single device on network 110 to facilitate simultaneous delivery of requests to both servers 102,104, as described below.

**[0055]** In steps 202A-B, the request is parsed by servers 102,104 to identify the requested asset. In steps 204A-B, the requested asset is retrieved from storage. In steps 206A-B, servers 102,104 begin processing the retrieved asset into video frames suitable for viewing at the client location.

**[0056]** In a preferred embodiment, the operational status of server 102 is verified on a continuing basis using one or more of the techniques described above. During each interval that no failure is detected (step 208=No), the video frame generated by primary server 102 for that interval is streamed to the client (step 210), and an identical video frame generated by secondary server 104 is discarded (step 212). In a preferred embodiment, this may be achieved by evaluating a "inhibit transmission" flag in the secondary server's network interface logic. When the flag is set, all packet transmissions are discarded, whereas when the flag is cleared the transmissions proceed. In step 214, unless the entire asset has been streamed (step 220), processing returns to steps 206A-B and the next video frame is prepared.

**[0057]** By contrast, if a failure is detected (step 208=Yes), a failover is triggered and the video frame generated by secondary server 104 is instead streamed to the client. In step 218, unless the entire asset has been streamed (step 220), processing proceeds to step 206B where the next video frame is prepared by the now-primary server 104.

**[0058]** In a preferred embodiment, the steps of verifying the primary server's operational status and, if necessary, initiating a failover to the secondary server are performed in less than one interval. Consequently, the failover is transparent to the client thus ensuring a satisfactory viewing experience.

**[0059]** Figs. 3A-3E are composite block/flow diagrams emphasizing the timing of data delivery in a failover pair and the zero-interval failover capability of the present system

and method. An interval, in the following description, is assumed to be equal to one NTSC video frame.

[0060] As shown in Fig. 3A, incoming requests are delivered to both primary server 102 and secondary server 104 in failover pair 100.

[0061] As shown in Fig. 3B, both the primary server and the secondary server respond by processing the request and preparing the first video frame (duration = one interval) for delivery (302A-B).

[0062] Fig. 3C illustrates that, at the end of interval 1, the primary server has transmitted the first data segment (304A), and the first data segment generated by the secondary server has been discarded (304B). In addition, both servers have prepared segment 2 (302A-B). As illustrated in Fig. 3D, the system repeatedly iterates through these steps, with the primary server transmitting each processed video frame and the secondary server discarding it.

[0063] For purposes of the present example, it is assumed that a failure occurs in the primary server during the interval corresponding to data segment 4. Accordingly, a failover is initiated as shown in Fig. 3E. In this example, the secondary server has automatically been re-designated the primary server in the failover pair, and the original primary server has been re-designated as the secondary server. The new primary server (former secondary server) preferably delivers segment 4 during the current interval. Accordingly, despite the failure, no data is lost in transit to the client and the client does not directly or indirectly perceive the failure.

[0064] Fig. 4 is a block diagram illustrating a preferred network configuration of public and private interfaces for supporting a failover pair. Shown in Fig. 4 are gigabit interfaces 412 for transmitting streaming data to users, a network switch 402 and a hub 404. Hub 404 is preferably connected to the network switch via a network connection 416. Servers 102,104 are preferably connected over a public network to hub 404 via network connections 414. Additionally, primary server 102 and secondary server 104 are preferably connected via a private network connection 110.

[0065] The public interface on each server preferably has two IP addresses associated with it, an administration IP address (admin\_IP) and a Stream IP address (stream\_IP). The

administration IP address is preferably unique across all servers on the network, while the stream IP address on each of the servers in a failover pair is preferably the same. This allows both the primary and secondary servers to see all control and data requests from clients. The private interface of each server is preferably assigned its own private IP address to facilitate communication over the private network.

[0066] In a preferred embodiment, each server preferably maintains two failover variables that define the server's state with respect to failover functionality: failover\_type and failover\_state. The failover\_type variable may preferably take the following values:

- FOTYPE\_MASTER - indicating that the server's current designation is as a primary server;
- FOTYPE\_SLAVE - indicating that the server's current designation is as a secondary server;
- FOTYPE\_UNDEF - indicating that the server is not currently designated as either a primary or a secondary server.

[0067] The failover\_state variable can preferably take the following values:

- FOSTATE\_LIVE - indicating that a partner exists and is online;
- FOSTATE\_UNDEF - indicating that the failover state is currently undefined;
- FOSTATE\_FAILED - indicating that a particular server is not operational.
- FOSTATE\_SYNCING - indicating that a synchronization is in process;
- FOSTATE\_SYNCED - indicating that synchronization is complete;
- FOSTATE\_NOPARTNER - indicating that a server has no failover partner;
- FOSTATE\_NONE - indicating that a particular server is offline.

[0068] In a preferred embodiment, a failover state is defined for each server that is a function of its failover\_type and failover\_state values. More specifically, the relationship between a server's failover state and its failover variables is preferably the following:

[0069] **failover state = Primary:**  
failover\_type = FOTYPE\_MASTER; and  
failover\_state = FOSTATE\_LIVE.

[0070] **failover state = Primary\_Offline:**  
failover\_type = FOTYPE\_MASTER; and  
failover\_state = FOSTATE\_NONE.

[0071]        **failover state = Primary\_No\_Secondary:**  
failover\_type = FOTYPE\_MASTER; and  
failover\_state = FOSTATE\_NOPARTNER.

[0072]        **failover state = Failed:**  
failover\_type = FOTYPE\_MASTER; and  
failover\_state = FOSTATE\_FAILED.

[0073]        **failover state = Secondary:**  
failover\_type = FOTYPE\_SLAVE; and  
failover\_state = FOSTATE\_LIVE.

[0074]        **failover state = Secondary\_Offline:**  
failover\_type = FOTYPE\_SLAVE; and  
failover\_state = FOSTATE\_NONE.

[0075]        **failover state = Secondary\_Synchronizing:**  
failover\_type = FOTYPE\_SLAVE; and  
failover\_state = FOSTATE\_SYNCING.

[0076]        **failover state = Secondary\_Synchronized:**  
failover\_type = FOTYPE\_SLAVE; and  
failover\_state = FOSTATE\_SYNCED.

[0077]        **failover state = Secondary\_No\_Primary:**  
failover\_type = FOTYPE\_SLAVE; and  
failover\_state = FOSTATE\_NOPARTNER.

[0078]        Fig. 5 is a state diagram illustrating these states and their relationship in a preferred embodiment. Shaded states indicate that at least one server is online (500A, 500B, 540A and 540B), while unshaded states indicate that no server is available (580, 582, 584, 586, 588). As described below, a failover state may change as a result of various events including the occurrence of a failover or the configuration, installation, or initialization of a server.

[0079]        A preferred embodiment for configuring a new failover pair 100 will now be described in conjunction with Fig. 5. To begin, an administrator preferably designates one server (e.g., server 102) as a primary server using a defined set\_config\_type command. In addition, the administrator preferably sets both the heartbeat frequency (i.e., the frequency at which heartbeat signals will be transmitted by the primary server) and the maximum number of missed heartbeats allowed before a failure is declared. Upon completion of these configuration steps primary server 102 transitions to Primary\_Offline state 580.

[0080] A start streaming command is preferably defined for bringing a server online. Since (in the present example) the secondary server has not yet been configured, when the administrator issues this command for primary server 102, it transitions to Primary\_No\_Secondary state 500B. At this point, primary server 102 can process requests and stream data to clients. No failover capability is available, however, since a failover partner for primary server 102 has not yet been configured.

[0081] The administrator next designates a second server (e.g., server 104) as a secondary server using the set\_config\_type command and specifies the secondary server's failover partner (e.g., by specifying the primary server's administration IP address and private IP address). Upon completion of these configuration steps secondary server 104 transitions to Secondary\_Offline 584.

[0082] The administrator next issues a failover\_sync command to synchronize the secondary server's failover-sensitive parameters with those of the primary server. The failover\_sync command causes secondary server 104 to transition to Secondary\_Synchronizing state 586. Successful synchronization causes the secondary server to transition to Secondary\_Synchronized state 588. If the synchronization fails, the secondary server's failover state returns to Secondary\_Offline 584.

[0083] The above-described synchronization preferably synchronizes a plurality of servers 102,104's parameters. Illustratively, these may include:

- System time
- Broadcast table (the list of currently active or scheduled sessions)
- Asset list
- Mount points (BASS)
- Stream\_Limits
  - MPEG\_Stream\_Count\_Limit (the maximum number of MPEG streams)
  - MPEG\_Bandwidth\_Limit\_In\_KBPS (the maximum bandwidth allotted to MPEG streams)
  - Server\_Bandwidth\_Limit\_In\_KBPS (the maximum total bandwidth allotted)
- Ports
  - MPEGRTSP
- Failover
  - Failover\_Heartbeat\_Freq (in milliseconds)
  - Max\_Missed\_Heartbeats
- Load\_Balance (see below)
  - Load\_Balance\_Active

Load\_Balance\_Group\_ID

- Network
  - StreamIP
  - Stream\_Mask
  - Routes

[0084] In a preferred embodiment, these synchronization tasks are performed by the secondary server. Once the secondary server is in a synchronized secondary failover state (e.g., Secondary\_Synchronized, Secondary\_No\_Primary, or Secondary), however, all updates to operational parameters are preferably handled via updates to the primary server. Any configuration changes made while the failover system is online are preferably synchronized across the failover partners via private network 110.

[0085] A failover\_unsync command is preferably defined for transitioning the secondary server from Secondary\_Synchronized state 588 to an offline state to enable administrative updates to server parameters.

[0086] Once the secondary server is in the Secondary\_Synchronized state, the start\_streaming command is used to bring the secondary server online. If no primary server is online, the secondary server transitions to the Secondary\_No\_Primary state 540B. In the present example, however, the primary server is already online when the secondary server receives the start\_streaming command. Accordingly, the secondary server transitions to Secondary state 540A. Concurrently, the primary server transitions from Primary\_No\_Secondary 500B to Primary 500A.

[0087] Once both primary and secondary servers are online, a final synchronization preferably takes place to establish communication channels between the servers via the private network. These communication channels are used to ensure synchronization of network activity between primary and secondary servers. For example, the primary server may utilize this channel during system operation to transmit stream IDs, packet numbers, and packet send times, and to ensure synchronization with the secondary server. Furthermore, the primary server may use this channel to transmit randomly assigned UDP ports to ensure that the primary and secondary servers use the same port numbers.

[0088] In a preferred embodiment, a stop\_streaming command is defined for taking a server offline. If a primary server goes offline at any time, its secondary server 104 transitions from Secondary state 540A to Secondary\_No\_Primary state 540B. Conversely, if

a primary server comes online, its secondary server transitions from Secondary\_No\_Primary state 540B to Secondary state 540A. Similarly, if a secondary server goes offline at any time, its primary server 102 transitions from Primary state 500A to Primary\_No\_Secondary state 500B. Conversely, if a secondary server comes online, its primary server's state transitions from Primary\_No\_Secondary state 500B to Primary state 500A.

**[0089]** It should be noted that the present system and method places no restriction on the order in which primary and secondary servers are brought online. If the primary server is brought online first, it will maintain a Primary\_No\_Secondary state until its secondary failover partner is brought online. Similarly, if the secondary server is brought online first, it will maintain a Secondary\_No\_Primary state until its primary failover partner is brought online.

**[0090]** It should further be noted that, in the example described above where the primary server is brought online before the secondary server and consequently may begin streaming with a failover state of Primary\_No\_Secondary, any VOD streams from the primary server before the secondary server comes online will not be backed up by the secondary server. However, broadcast sessions running before the secondary server comes online are preferably backed up, since they will be started on the secondary server when it comes online based on the broadcast table, which is synchronized with the primary server's broadcast table.

**[0091]** A preferred embodiment for conducting a failover in the present system and method is now described. For purposes of illustration, it is assumed in the following description that a failover pair has been configured and brought online and that primary server 102 is in Primary state 500A and secondary server 104 is in Secondary state 540A.

**[0092]** As noted above, primary server 102 is preferably continuously monitored in one or more ways to permit immediate detection of server failure. For purposes of illustration it is assumed in the following description that both local monitoring by a server and heartbeat monitoring by the secondary server are implemented.

**[0093]** In this preferred embodiment, each server is provided with a kernel responsible for managing its server's failover\_type and failover\_state variables. Variable values may be set by the application layer (during, e.g., server configuration by the administrator) or by the kernel itself (e.g., when a failure is detected, as described below). In

a preferred embodiment, whenever a failover\_type or failover\_state value changes, the kernel transmits a signal to all application processes to inform them of the change (SIGFAILOVER). Application processes may query these variables at any time, and preferably check them upon receiving a SIGFAILOVER signal. As noted, the value of these two variables are tightly coupled with the application level set\_config\_type designation (i.e., Primary/Secondary designation) and failover state. Mappings for these couplings are described above.

[0094] In a preferred embodiment, local monitoring by the primary server is facilitated by providing a registration feature for critical processes which allows an application to identify itself to the kernel as a critical process. As shown in Fig. 5, when a registered critical process fails (e.g., application core, kernel trap, etc.), the kernel preferably transitions the primary server to Failed state 582, and sends a prebuilt failover message to the secondary server. To accomplish this state transition, the primary server's failover\_type is set to FOTYPE\_FAILED. The kernel also preferably disables the primary server's processing and generation of control messages and data transmission to all clients.

[0095] In a preferred embodiment, a failover message from the primary server may be the result of any number of conditions. Illustratively, these may include:

- Public Interface Failure
- Gigabit Interface Failure
- Disk Failure
- System Software Failure
- System Temperature Exceeds Operational Limits

[0096] In addition, the primary server's kernel is preferably adapted to send heartbeats at the configured frequency (Failover\_Heartbeat\_Freq) to the secondary server. The secondary server's kernel receives these heartbeats, expecting them at the same frequency because the systems are synchronized. If the secondary server misses more than the configured allowed heartbeats (Max\_Missed\_Heartbeats), the kernel initiates a failover, as described below.

[0097] A missed-heartbeat failure typically occurs when a failure is so catastrophic that the primary server is unable to transmit a failover message. Illustratively, the loss of a heartbeat may be the result of any of the following conditions:

- Private Interface Failure

- Power Supply Failure
- System Wide Failure
- Server Crash/Lockup (BSOD)

[0098] In a preferred embodiment, when a secondary server's kernel detects a failure by either missed heartbeats or by receiving a failover message from the primary server, the kernel initiates a failover by changing its failover type variable from Secondary to Primary and sending a SIGFAILOVER signal to some or all application processes preferably including all applications that manage client connections and stream data and those handling configuration synchronization. As a result of the change to its failover type variable and the transition of primary server 102 to the Failed state, the secondary server transitions to the Primary\_No\_Secondary state, as shown in Fig. 5. In a preferred embodiment, the secondary server also transmits a message to the primary server, if necessary, to instruct it to cease streaming data to the client. If possible, processes on the primary server attempt to log the failure and then go into an idle mode.

[0099] As will be recognized, to ensure exact mirroring of the primary server, the secondary server preferably is not considered a separate available server in load balancing determinations such as those described in U.S. patent application serial No. \_\_\_\_\_, filed \_\_\_\_\_, entitled "System and Method For Digital Media Serve Load Balancing," which is hereby incorporated by reference in its entirety for each of its teachings and embodiments. Rather, the secondary server mirrors any load balance tasks that are currently assigned to the primary server. In other words, from a load balancing perspective, the primary server and secondary servers are treated as one server. Preferably, however, the secondary server does not transmit load balance data, i.e., the primary server is preferably responsible for sending out load information. It should also be noted that where the load balance (e.g., where the algorithm randomly assigns a task in the case of a load tie) and the load balancer accordingly assigns new sessions randomly, tasks randomly allocated to the primary server are preferably communicated between the primary and secondary servers to ensure identical session maintenance.

[00100] While the invention has been described in connection with specific embodiments, it is evident that numerous alternatives, modifications, and variations will be apparent to those skilled in the art in light of the foregoing description.